Proc. Natl. Acad. Sci. USA
vVol. 87, pp. 4236–4240, June 1990
Biochemistry

PRP5: A helicase-like protein required for mRNA splicing in yeast

(RNA helicase/yeast *PRP* genes/pre-mRNA splicing)

GLORIA DALBADIE-MCFARLAND AND JOHN ABELSON

Division of Biology, California Institute of Technology, Pasadena, CA 91125

Contributed by John Abelson, March 5, 1990

ABSTRACT **A 96-kDa protein predicted by the DNA sequence of the *Saccharomyces cerevisiae PRP5* gene contains a domain that bears a striking resemblance to a family of RNA helicases characterized by the conserved amino acid sequence Asp-Glu-Ala-Asp (D-E-A-D). Previous work indicated that the product of the *PRP5* gene is required for splicing and that spliceosome assembly does not occur in its absence. However, its precise role in splicing and the nature of its biochemical activity remained unknown. To examine the role of *PRP5* in splicing, we cloned the gene by complementation of a temperature-sensitive mutation and determined its DNA sequence. We discuss here the possible roles for an RNA helicase in splicing and for the activity of the PRP5 protein.**

Splicing of introns from nuclear pre-mRNAs occurs by a mechanism that appears to be common to most eukaryotes and requires a complex trans-acting apparatus. Small nuclear ribonucleoproteins (snRNPs) U1, U2, U4/U6, and U5 and a number of proteins assemble with a pre-mRNA substrate into a large complex known as the spliceosome (reviewed in refs. 1–3). This assembly process and the subsequent excision of the intron in the form of a lariat and ligation of the exons are ATP dependent. Small nuclear RNAs (snRNAs) in the spliceosome are thought to have a role in catalysis and in splice-site recognition and alignment. A number of the protein factors as well as the snRNA components of the spliceosome are essential for nuclear splicing. The nuclear splicing mechanism differs in this regard from the self-splicing mechanisms it may have evolved from, which require cis-acting catalytic RNA elements but no essential protein factors. Understanding the role of proteins in splicing, therefore, not only will provide better understanding of the nuclear splicing mechanism itself but also may throw light on the pathway of evolution from self-splicing to splicing mediated by a trans-acting apparatus.

Identification of a number of proteins essential for nuclear splicing has been greatly facilitated in the yeast *Saccharomyces cerevisiae* by the isolation of temperature-sensitive (ts) mutants that affect pre-mRNA splicing. These mutants, originally called *rna* mutants, have been renamed *prp* mutants, for pre-RNA processing. *In vitro* biochemical studies (4, 5) of the original set of mutants, *prp2* to *-11* (6), showed that their gene products are directly involved in splicing. Of these, the products of *PRP3*, *-4*, *-5*, *-7*, *-8*, and *-11* are required for spliceosome assembly (5), whereas the *PRP2* product acts after a spliceosome forms but before any covalent changes occur on the pre-mRNA substrate. Additional *prp* mutants recently isolated (7) have novel phenotypes which affect other late stages in splicing, including 3′ splice site cleavage and exon ligation, as well as degradation of the lariat intron. Many of the *PRP* genes have been cloned (8–11), and in all known cases they encode proteins. *PRP4* and *PRP8* are known to be snRNP proteins specifically associated with U4/U6 and U5, respectively (12–14). Although the step(s) of the splicing pathway affected by these proteins are known, and some significant molecular interactions have been identified, the nature of the biochemical activity associated with these proteins is not understood.

We report here the finding that PRP5,* a protein directly involved in nuclear splicing and spliceosome assembly, has a central domain that is strikingly similar to the "D-E-A-D family" RNA helicases, which share the sequence Asp-Glu-Ala-Asp (D-E-A-D). Recently defined on the basis of amino acid sequence similarity and shared biochemical properties, the RNA helicase family includes proteins found in organisms ranging from *Escherichia coli* to humans, among them the eukaryotic translation initiation factor 4A (eIF4A) (15).

MATERIALS AND METHODS

Strains. *S. cerevisiae* strains SPJ 5.41 (α, *his7*, *leu2*, *ura3-52*, *prp5*) and A364A (**a**, *ade1*, *ade2*, *lys2*, *tyr1*, *his7*, *ura1*, *gal1*) were obtained from A. Lustig (Sloan–Kettering Cancer Center, New York). *S. cerevisiae* 5.41::pGD561 (α, *his7*, *leu2*) was constructed by chromosomal integration of plasmid pGD561, linearized with *Sac* I, into strain SPJ 5.41. Transformation of yeast strains was by the lithium acetate method (16).

E. coli strains used were DH5-α [φ80d*lacZ*Δ M15, *endA1*, *recA1*, *hsdR17*, *supE44*, *thi-1*, *gyrA*, *relA1*, Δ(*lacZYA–argF*) U169] and DH5-α F′ [F′, φ80d*lacZ*ΔM15, *endA1*, *recA1*, *hsdR17*, *supE44*, *thi-1*, *gyrA*, *relA1*, Δ(*lacZYA–argF*)U169], obtained from Bethesda Research Laboratories.

Plasmids and Library. YCp50 is a centromere-containing yeast–*E. coli* shuttle vector derived from pBR322, which confers ampicillin resistance to *E. coli* and carries the yeast *URA3*, *CEN4*, and *ARS1* genes. pSEYC68, obtained from S. Emr (California Institute of Technology, Pasadena, CA), is similar to YCp50 and contains, in addition, a copy of the *E. coli lacZ* gene interrupted by a polylinker containing multiple unique cloning sites. pGD551 is the original *PRP5* isolate consisting of a 9-kilobase (kb) insert cloned in the YCp50 *Bam*HI site. YIp5 is a yeast integrating vector that carries the yeast *URA3* gene and confers ampicillin resistance to *E. coli* (17). pGD561 consists of a 5.6-kb *Pvu* II–*Cla* I fragment that complements the *prp5* ts defect, cloned in YIp5 between unique *Nru* I and *Cla* I sites. The yeast genomic library screened to isolate the *PRP5* gene, a gift from E. Phizicky (University of Rochester, Rochester, NY), had been constructed by ligating yeast genomic DNA partially digested with *Sau*3A into the *Bam*HI site of YCp50.

Cloning. A yeast genomic library, constructed in a centromere-containing vector which carries the yeast *URA3* gene and confers ampicillin resistance to *E. coli*, was used to

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: snRNP, small nuclear ribonucleoprotein; snRNA, small nuclear RNA; eIF4A, eukaryotic translation initiation factor 4A; ts, temperature-sensitive.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M33191).

transform a *prp5* ts strain, *S. cerevisiae* SPJ 5.41 by the lithium acetate method (16). *URA*$^+$ transformants were selected on agar plates of synthetic defined medium lacking uracil, incubated at a permissive temperature (23°C). These transformants were replica plated onto fresh plates and incubated at a nonpermissive temperature (35°C) to select temperature-resistant colonies. DNA was prepared from resistant colonies and used to transform *E. coli* to ampicillin resistance.

DNA Sequence Determination and Analysis. The sequence of *PRP5* was determined by the dideoxynucleotide method (18). A 3.0-kb *Hin*cII–*Cla* I fragment capable of complementing the *PRP5* ts defect was sequenced in both strands. A 3.9-kb *Eco*RI fragment cloned in M13 mp18 containing the entire *Hin*cII–*Cla* I fragment was sequenced in one orientation, using specific oligonucleotide primers (a to i). The primers, derived from the sequence, were synthesized sequentially as it was determined (19). The sequence of the opposite strand was determined on a double-stranded template (pGD561), also using synthetic oligonucleotide primers derived from the sequence (a′ to i′). Sequence data were assembled and analyzed initially using programs of the BIONET National Computer Resource for Molecular Biology. The assembly and analysis were completed by using the University of Wisconsin Genetics Computer Group (UWGCG) Sequence Analysis Software Package, version 6.1 (copyright 1989 by John R. Devereux), run on a VAXstation 2000 computer. Protein sequence data bases were searched with the program FASTA (20); eIF4A was the highest-scoring sequence in the National Biomedical Research Foundation data base (release 22 of September 1989) with PRP5 as the query sequence. The GenBank DNA sequence data base (release 62.0 of December 1989) was searched with the TFASTA (20) program, which translates the DNA sequences in the data base in all six reading frames. Sequences not present in these databases were compared to PRP5 by using the UWGCG program BESTFIT and multiple sequences were aligned with LINEUP.

Oligonucleotides. The following synthetic oligonucleotides, shown schematically in Fig. 1, were used as DNA sequencing primers: a, 5′-TGAGACTGTGGACGATTC; b, 5′-ATAGGGAGTCTTTATTG; c, 5′-TGACGTCATTGAAGGAG; d, 5′-GACTTAGTCTGGATAAT; e, 5′-ACGGCCATTATCAAACA; f, 5′-GAGATTGTTGTTGCCAC; g, 5′-GTTCGATTTAGGTTTTG; h, 5′-GTGTCGTTGGTTATTATC; i, 5′-GTTTATATTAATGATTT; a′, 5′-TCTACATCATATATTAT; b′, 5′-CATGCTTGATAAATAATAAGGTTG; c′, 5′-CTCTGGCAGTTCTTCCAGTAGTAT; d′, 5′-CTTCAACATCGCTTGAC; e′, 5′-TGTTCGGAAAAGTTGCACTAAATAG; f′, 5′-CTTCTGTGAATTTAGTA; g′, 5′-TCCGTTGATAGTCCCAG; h′, 5′-TCGTCGGTTCCTCCAAG; and i′, 5′-ACGTTCCTGTCTTTCCG.

DNA Blot Analysis. Genomic yeast DNA was completely digested with *Bst*EII and *Sal* I and analyzed on 0.7% agarose gels. DNA was transferred to a GeneScreen membrane (NEN) according to the manufacturer's instructions. Probes were prepared by radiolabeling gel-purified DNA fragments by the random primer method (21). Prehybridization was done in 4× SSC (1× SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 1× Denhardt's solution (0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin) at 67°C for 6 hr. Hybridization was done in 4× SSC/1× Denhardt's solution/0.1 mg of salmon sperm DNA per ml/0.2% SDS containing ^{32}P-labeled probe, for 18 hr at 67°C. Blots were washed twice at room temperature in 2× SSC and once at 41°C in 0.1× SSC.

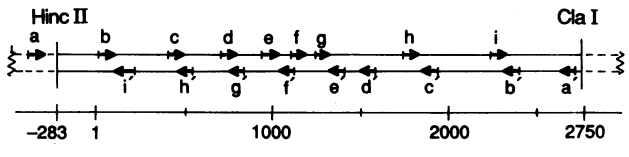

FIG. 1. Schematic representation of the complementing *Hin*cII–*Cla* I fragment, showing the relative positions of the oligodeoxynucleotides used for sequencing. The arrows represent oligonucleotides, pointing in a 5′ to 3′ direction. Nucleotide number 1 in the sequence is the A in the first ATG in a 96-kDa open reading frame.

RESULTS

Cloning, Complementation Analysis, and Subcloning. A genomic DNA fragment containing *PRP5* was isolated from a yeast genomic library by complementation of the ts defect of the yeast prp5 mutant. Of the 17,000 colonies screened, 3 complemented the *prp5* defect. All plasmids obtained contained the same 9-kb insert according to their restriction maps; one was selected for all subsequent studies and named pGD551. The 9-kb insert from pGD551 complements the *prp5* ts growth defect when reintroduced into *S. cerevisiae* SPJ 5.41. The position of the *PRP5* gene within this insert was mapped by subcloning a number of overlapping restriction fragments in centromere-containing vectors and testing for their ability to complement the *prp5* defect. The smallest complementing fragment is 3.0 kb (Fig. 2).

Genetic Mapping by Integrative Transformation. *Integration.* A *Pvu* II–*Cla* I 5.6-kb complementing fragment was subcloned in an integrating vector (YIp5). The resulting plasmid pGD561, containing both the *URA3* gene and the putative *PRP5* gene, was cut at a unique site within the 5.6-kb insert, to direct the integration to a homologous locus in the chromosome (22). A *ura3-52*, *prp5* strain, SPJ 5.41, was transformed with the linearized DNA, resulting in *URA*$^+$, temperature-resistant transformants. The site of integration was confirmed by DNA blot analysis (Fig. 3), using radiolabeled 5.6-kb *Pvu* II–*Cla* I fragment as a probe. A single band was observed in lanes containing DNA from wild type or from SPJ 5.41 strains, and two bands, predicted from the experimental design (22), were observed in those lanes containing DNA from the transformed strain. The results indicate that a single copy of the plasmid was integrated at a homologous locus and that the cloned DNA represents a unique sequence in the *S. cerevisiae* genome. The latter conclusion was corroborated by other DNA blot analyses in which different restriction enzymes were used to digest genomic DNA (not shown).

Tetrad analysis. A diploid strain resulting from a cross between the transformed *S. cerevisiae* strain 5.41::pGD561

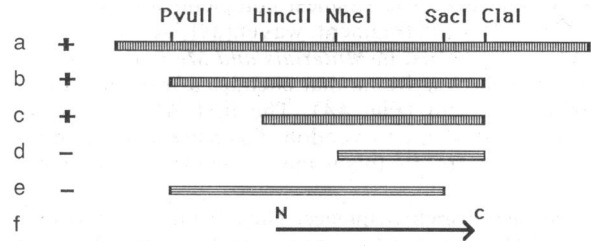

FIG. 2. Complementation analysis. The horizontal bars represent DNA fragments that were tested for their ability to complement a *prp5* ts defect and which define the position of the gene. The fragments, which were cloned in centromere-containing vectors, are a, the original 9-kb complementing fragment; b, a 5.6-kb *Pvu* II–*Cla* I fragment; c, a 3.0-kb *Hin*cII–*Cla* I fragment; d, a 1.7-kb *Nhe* I–*Cla* I fragment; and e, a 5.0-kb *Pvu* II–*Sac* I fragment. The arrow f represents the open reading frame found in the DNA sequence, positioned relative to the fragments; correspondence to the N and C termini of the protein is indicated. Fragments a, b, c, and e were cloned in YCp50; fragment d was cloned in pSEYC68. Fragments a, b, and c complement, whereas fragments d and e do not.

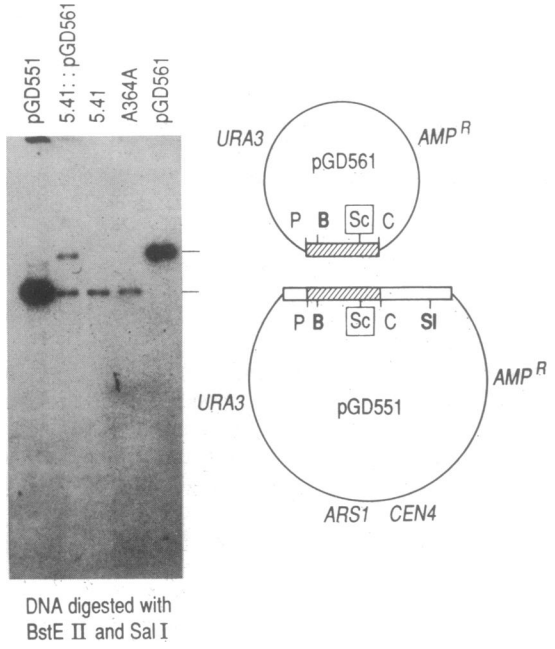

FIG. 3. A blot containing genomic DNA from a wild-type yeast strain, from strain 5.41, and from strain 5.41 transformed with the linearized pGD561 was probed with the radiolabeled 5.6-kb *Pvu* II–*Cla* I complementing fragment. Restriction sites B (*Bst*EII), C (*Cla* I), P (*Pvu* II), and Sc (*Sac* I), the integration site, are present in pGD551, pGD561, and the chromosome; SI, (*Sal* I) is present in pGD551 and in the chromosome but not in the integrating plasmid pGD561.

and a wild-type strain, A364A, was sporulated. By tetrad analysis (15 tetrads), the wild-type, temperature-resistant phenotype segregated 4:0, the URA3$^+$ phenotype 3:1, and all other markers were found to segregate as predicted. The analysis showed that the plasmid was integrated at the *PRP5* locus or at a closely linked site rather than at an unlinked suppressor site. This linkage strongly suggests that the cloned DNA is indeed the *PRP5* gene. The 3:1 segregation of the URA3$^+$ phenotype is due to the presence of *URA3* at two different loci in the parental strains used in the cross: the *URA3* locus in A364A and the site of the integration event in 5.41::pGD561. Crosses of strains combining a wild-type PRP5 phenotype with a ura3 phenotype to 5.41::pGD561 should result in diploids carrying the *URA3* marker only at the site of integration; however, the diploids obtained did not sporulate well, and so they were not used for this analysis.

DNA Sequence. The smallest complementing fragment, a 3.0-kb *Hin*cII–*Cla* I fragment, was entirely sequenced in both strands as described in *Materials and Methods*.

An open reading frame that encodes a 96-kb hypothetical protein is present (Fig. 4*A*). The first AUG in the open reading frame is only two codons downstream of an in-frame stop codon; therefore this is most likely the initiation codon used.

Translated Protein Sequence. The most noteworthy feature of the PRP5 protein sequence is a striking resemblance to a family of RNA helicases known as the D-E-A-D family (Fig. 4*B*) (15). The helicase most closely resembling PRP5 is p68, a human nuclear protein (23, 24). PRP5 and p68 are 40% identical and 63% similar in the region of similarity, which spans approximately 400 residues of PRP5. Twenty-two of the 25 residues that are absolutely conserved in all known RNA helicases are also present in PRP5 at conserved relative positions (bold in Fig. 4). This 400-residue region of similarity is a central segment of PRP5 (domain 2) flanked by 260 amino-terminal residues (domain 1) and 189 carboxyl-terminal residues (domain 3) of unique sequence. These may represent physically distinct domains unique to PRP5. Domains 1 and 3 are more highly charged than the helicase domain, with an aspartate + glutamate + histidine + lysine + arginine content of 42% in domain 1 and 37% in domain 3, whereas it is 26% in the helicase domain. Relatively high contents of lysine and glutamate are found in domains 1 (14.8% lysine, 12.2% glutamate) and 3 (14.6% lysine, 11.9% glutamate), approximately twice that of the helicase domain (7.2% lysine, 6.7% glutamate). There is a potential nuclear localization signal (25) at position 90 of the protein sequence, in the amino-terminal unique sequence domain. Sequences possibly related to RNA-binding protein motifs (26) occur in domain 3 at positions 679 and 790. Other motifs present in nucleic acid binding proteins, including the zinc finger consensus (27), leucine zipper motif (28), and tRNA synthetase signature sequence (29), are not present in PRP5. No additional significant similarities have been found.

DISCUSSION

We have cloned a yeast DNA fragment that complements a *prp5* ts growth defect and is linked to the *PRP5* locus. The cloned DNA therefore probably is the *PRP5* gene. Its DNA sequence contains an open reading frame that encodes a 96-kDa hypothetical protein, which we will refer to as PRP5 hereafter.

PRP5 closely resembles members of a recently identified family of RNA helicases characterized by several invariant sequence motifs (15, 30). The characteristic biochemical activity, ATP-dependent RNA unwinding, has been demonstrated experimentally in a few of these proteins (31–33). PRP5 and probably SPP81 are involved in nuclear splicing in yeast. SPP81, a suppressor of the *prp8* mutation, is also a putative helicase (D. Jamieson and J. Beggs, personal communication). Other putative helicases implicated in RNA processing are MSS116 (34), which is involved in yeast mitochondrial splicing, and SPB4, which is involved in rRNA processing (35).

Special forms of ATP-binding proteins' A (D-X_4-A-X_4-G-K-T) and B (D-E-A-D) motifs are among the highly conserved motifs shared by RNA helicases (36–38). The functions of various other conserved residues and motifs found at similar relative positions in all the proteins are not known. PRP5 shares most of the regions of amino acid conservation found in other RNA helicases, and it has unique sequence segments at both the amino and carboxyl termini, as do most other helicases whose entire sequence is known (Fig. 4*B*) (15). These unique domains could conceivably tailor the helicases' activity for their specific functions, providing cellular localization signals and perhaps specific interaction sites.

Three differences between PRP5 and other helicases are worth noting. (*i*) An alanine residue at position 6 of the A motif is thought to be characteristic of RNA helicases, as distinct from other classes of ATP-binding proteins, in which a glycine residue is found at that position. The alanine residue occurs in all known members of the helicase family, except PRP5 and SPB4, where it is replaced by a serine, which is similar in size to alanine although more hydrophilic. (*ii*) The residues between positions 470 and 495 in the PRP5 translated sequence represent a spacer region between conserved regions that diverges in length: it is 9–17 residues longer than similar regions in other known helicases. This may reflect a difference in function, or perhaps it marks the position of the surface of the protein where differences in primary sequence are tolerated more readily without disruption of tertiary structure (39). (*iii*) The sequence **Y X H *R I* G R**, conserved in 13 of 15 helicases known to date, becomes **Y X H *T T* G R** in PRP5, where a charged residue (R) and a hydrophobic residue (I), are replaced by two smaller, moderately hydrophilic, residues (T T). The sequence of this conserved region

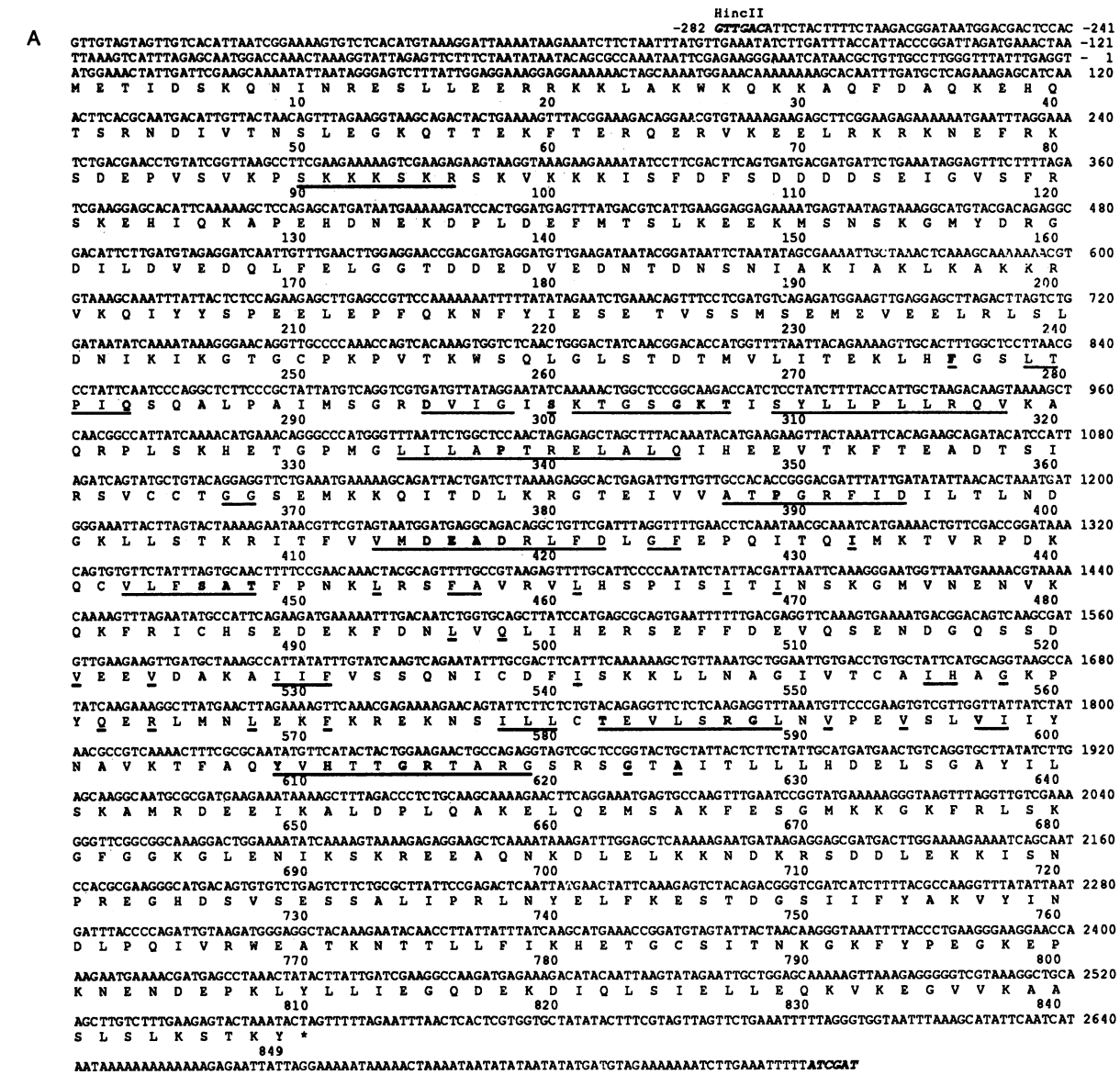

FIG. 4. (*A*) Nucleotide sequence of the *Hin*cII–*Cla* I fragment that complements the *prp5* ts defect. Translation is shown below in single-letter code. A possible nuclear localization signal at position 90 in the translated protein sequence is underlined. Regions of similarity with RNA helicase conserved residues are also underlined, and absolutely conserved positions appear in bold characters. The region of alignment with the RNA helicase family extends approximately from residue 260 to residue 630. (*B*) Alignment of PRP5 and members of the helicase family from different species. Conserved residues appear in bold characters.

also diverges in SPB4. The significance of these differences is not known. They could define functionally distinct members of the helicase family or simply be tolerated changes in noncrucial residues.

The characteristics of the PRP5 protein sequence suggest that PRP5 possesses a helicase activity responsible for the ATP-dependent unwinding of base-paired RNA. Splicing is ATP dependent, and a number of base-pairing interactions, such as those between U1 and the 5′ splice site, U2 and the branch point, or U4 and U6, must be formed or resolved during the course of a splicing reaction (2, 3). Likely, a helicase directly involved in splicing will participate by converting snRNAs between active and inactive forms, or by making regions of the pre-mRNA substrate and snRNAs available or unavailable for base pairing or covalent modification. Other PRP proteins have been shown to be integral snRNP proteins (12, 13). We have no convincing evidence to indicate that PRP5 is an integral snRNP protein, but the possibility cannot be rigorously ruled out at this point.

PRP5 acts early in the splicing pathway (2). U2 snRNP binding, the first step in splicing known to require ATP (40–43), requires functional PRP5, whereas U1 snRNP binding, which occurs prior to U2 binding whether ATP is present or not (44), does not require functional PRP5 (S. Ruby, personal communication). One hypothesis is that the putative PRP5 helicase activity is required for U1, the pre-mRNA substrate, or U2 to adopt conformations that allow U2 binding. Alternatively, it may be necessary for PRP5 to be incorporated into the spliceosomal complex prior to U2 binding, but it may carry out its activity later in the pathway. For example, it has been proposed that a highly conserved region of U6 is involved in catalysis but remains inactive while base paired to U4 (45). The U4/U6 interaction is destabilized before or during excision of the 5′ exon (40–42, 46), presumably freeing U6 to carry out its proposed catalytic function. The U4/U6 helix is stable (47) and is thought to require an energy-dependent unwinding step to release the proposed U6 catalytic activity. A helicase such as PRP5 or SPP81 could carry out this function.

We thank A. B. Sachs and R. Davis for communicating the SPB4 sequence prior to publication and D. Jamieson and J. Beggs for communicating unpublished results. We also thank S. Ruby, J. Arenas, and P. Fabrizio for discussions and critical reading of the manuscript and S. Horvath for oligonucleotide synthesis. G.D.-M. was supported by an American Cancer Society Fellowship. This work was also supported by National Institutes of Health Grant GM 32637 awarded to J.A.

1. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. & Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55,** 1119–1150.
2. Guthrie, C. & Patterson, B. (1988) *Annu. Rev. Genet.* **22,** 387–419.
3. Steitz, J. A., Black, D. L., Gerke, V., Parker, K. A., Krämer, A., Frendewey, D. & Keller, W. (1988) in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, ed. Birnstiel, M. L. (Springer, Heidelberg), 115–154.
4. Lustig, A. J., Lin, R.-J. & Abelson, J. (1986) *Cell* **47,** 953–963.
5. Lin, R.-J., Lustig, A. & Abelson, J. (1987) *Genes Dev.* **1,** 7–18.
6. Hartwell, L. H. (1967) *J. Bacteriol.* **93,** 1662–1670.
7. Vijayraghavan, U., Company, M. & Abelson, J. (1989) *Genes Dev.* **3,** 1206–1216.
8. Last, R. L., Stavenhagen, J. B. & Wooford, J. L. (1984) *Mol. Cell. Biol.* **4,** 2396–2405.
9. Lee, M. G., Young, R. A. & Beggs, J. D. (1984) *EMBO J.* **3,** 2825–2830.
10. Soltyk, A., Tropak, M. & Friesen, J. D. (1984) *J. Bacteriol.* **160,** 1093–1100.
11. Jackson, S. P., Lossky, M. & Beggs, J. D. (1988) *Mol. Cell. Biol.* **8,** 1067–1075.
12. Banroques, J. & Abelson, J. (1989) *Mol. Cell. Biol.* **9,** 3710–3719.
13. Bjørn, S. P., Soltyk, A., Beggs, J. D. & Friesen, J. D. (1989) *Mol. Cell. Biol.* **9,** 3698–3709.
14. Lossky, M., Anderson, G. J., Jackson, S. P. & Beggs, J. (1987) *Cell* **51,** 1019–1026.
15. Linder, P., Lasko, P., Ashburner, M., Leroy, P., Nielsen, P., Nishi, K., Schnier, J. & Slonimski, P. (1989) *Nature (London)* **337,** 121–122.
16. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983) *J. Bacteriol.* **153,** 163–168.
17. Botstein, D., Falco, C., Stewart, S. E., Brennan, M., Scherer, S., Stinchcomb, D., Struhl, K. & Davis, R. (1979) *Gene* **8,** 17–24.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Strauss, E. C., Kobori, J. A., Siu, G. & Hood, L. E. (1986) *Anal. Biochem.* **154,** 353–360.
20. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
21. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
22. Orr-Weaver, T., Szostak, J. & Rothstein, R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6354–6358.
23. Hirling, H., Scheffner, M., Restle, T. & Stahl, H. (1989) *Nature (London)* **339,** 562–564.
24. Ford, M. J., Anton, I. A. & Lane, D. P. (1988) *Nature (London)* **322,** 736–738.
25. Smith, A. E., Kalderon, D., Roberts, B. L., Colledge, W. H., Edge, M., Gillett, P., Markham, A., Paucha, E. & Richardson, W. D. (1985) *Proc. R. Soc. London Ser. B* **226,** 43–58.
26. Bandziulis, R. J., Swanson, M. S. & Dreyfuss, G. (1989) *Genes. Dev.* **3,** 431–437.
27. Berg, J. M. (1986) *Science* **232,** 485–487.
28. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. (1988) *Science* **240,** 1759–1764.
29. Schimmel, P. (1987) *Annu. Rev. Biochem.* **56,** 125–158.
30. Chang, T.-H., Arenas, J. & Abelson, J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 1571–1575.
31. Ray, B. K., Lawson, T. G., Kramer, J. C., Cladaras, M. H., Grifo, J. A., Abramson, R. D., Merrick, W. C. & Thatch, R. E. (1985) *J. Biol. Chem.* **260,** 7651–7658.
32. Abramson, R. D., Dever, T. E., Lawson, T. G., Ray, B. K., Thatch, R. E. & Merrick, W. C. (1987) *J. Biol. Chem.* **262,** 3826–3832.
33. Nishi, K., Morel-Diville, F., Hershey, J. W. B., Leighton, T. & Schnier, J. (1989) *Nature (London)* **336,** 496–498.
34. Séraphin, B., Simon, M., Boulet, A. & Faye, G. (1989) *Nature (London)* **337,** 84–87.
35. Sachs, A. B. & Davis, R. W. (1990) *Science* **247,** 1077–1079.
36. Hodgman, T. C. (1988) *Nature (London)* **333,** 22–23.
37. Hodgman, T. C. (1988) *Nature (London)* **333,** 578.
38. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) *EMBO J.* **1,** 945–951.
39. Reidhaar-Olson, J. F. & Sauer, R. T. (1988) *Science* **241,** 53–57.
40. Cheng, S.-C. & Abelson, J. (1987) *Genes Dev.* **1,** 1014–1027.
41. Pikielny, C. W., Rymond, B. C. & Rosbash, M. (1986) *Nature (London)* **324,** 341–345.
42. Lamond, A. I., Konarska, M. M., Grabowski, P. J. & Sharp, P. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 411–415.
43. Konarska, M. M. & Sharp, P. A. (1987) *Cell* **49,** 763–774.
44. Ruby, S. & Abelson, J. (1988) *Science* **242,** 1028–1035.
45. Brow, D. A. & Guthrie, C. (1989) *Nature (London)* **337,** 14–15.
46. Blencowe, B. J., Sproat, B. S., Ryder, U., Barabino, S. & Lamond, A. I. (1989) *Cell* **59,** 531–539.
47. Brow, D. A. & Guthrie, C. (1988) *Nature (London)* **334,** 213–218.